SAI NIKITH REDDY YARRAM (700742917)

MACHINE LEARNING

CRN: 30521

ASSIGNMENT 2

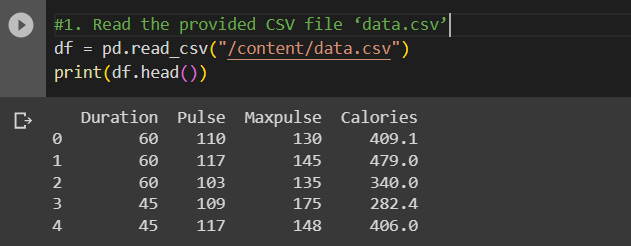
JUNE 22*nd*2023

GITHUB LINK: <https://github.com/nikithreddy30/MLAssignment2>

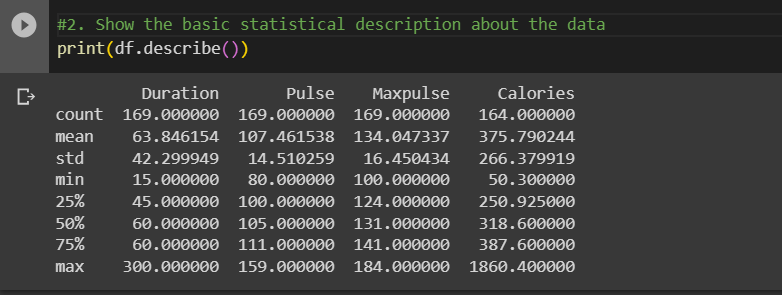
VIDEO LINK: <https://drive.google.com/file/d/1MW-Abr3wolZVLyTi0Q_LqGX6_ZQomYC1/view?usp=sharing>

**1. Pandas**

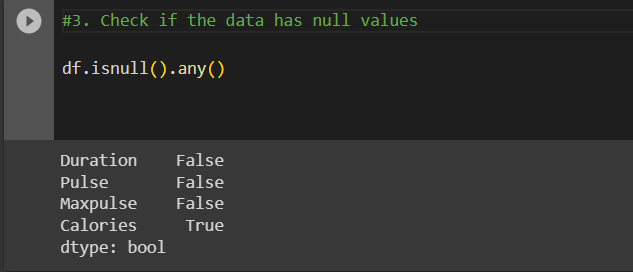
1. Read the provided CSV file ‘data.csv’ Print array shape.



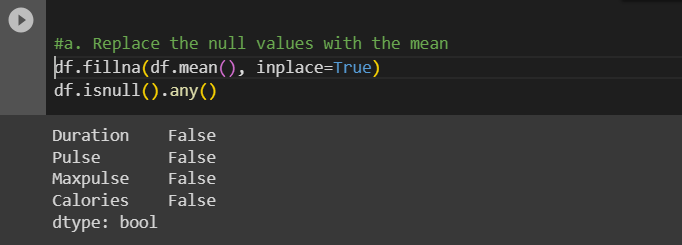
1. Show the basic statistical description about the data



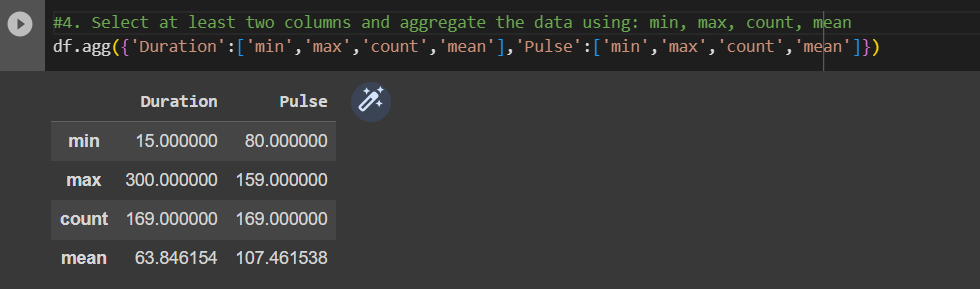
1. Check if the data has null values



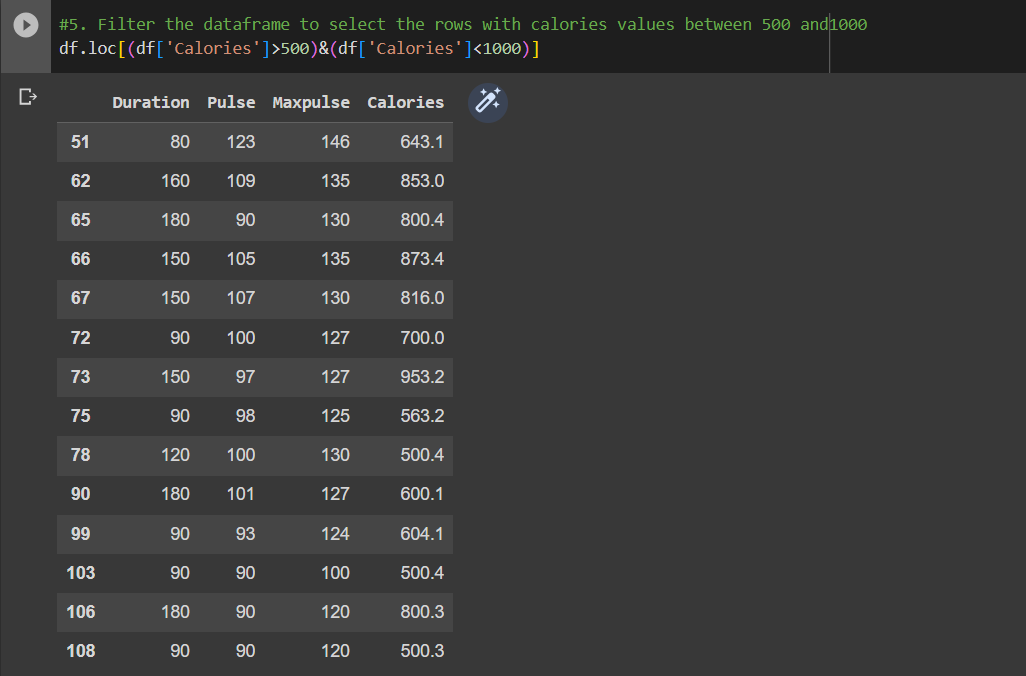
1. Replace the null values with the mean



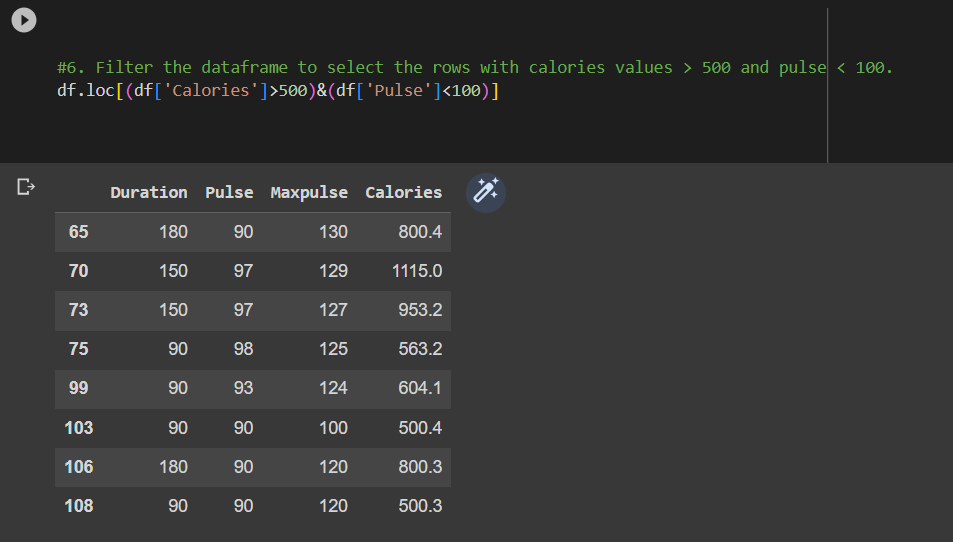
1. Select at least two columns and aggregate the data using: min, max, count, mean



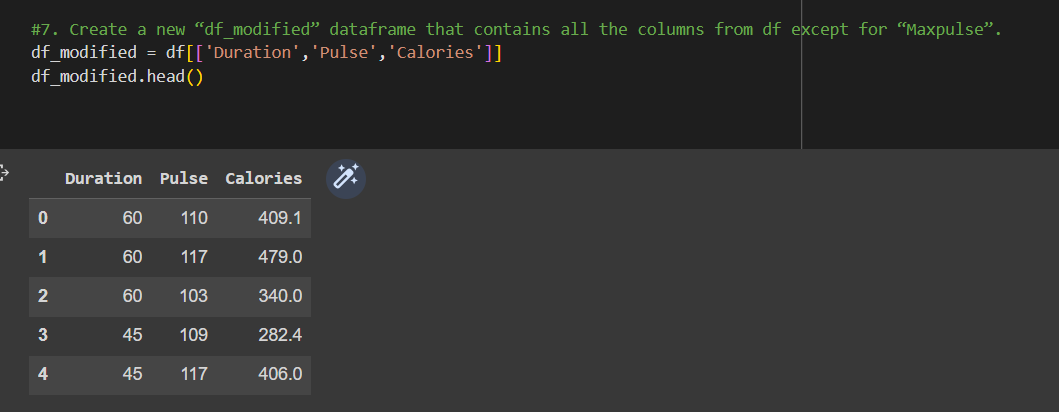
1. Filter the dataframe to select the rows with calories values between 500 and1000



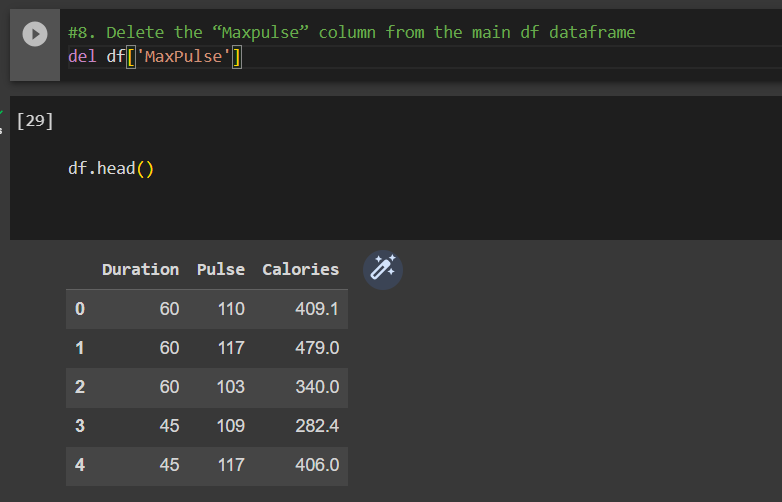
6.Filter the dataframe to select the rows with calories values > 500 and pulse <100



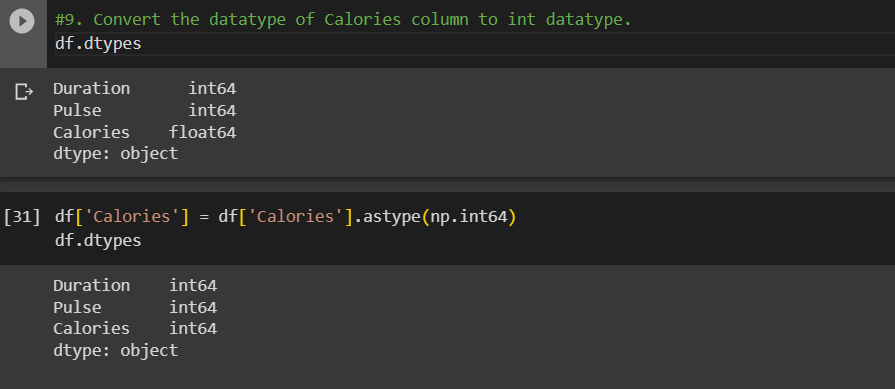
7. Create a new “df\_modified” dataframe that contains all the columns from df except for “Maxpulse”.



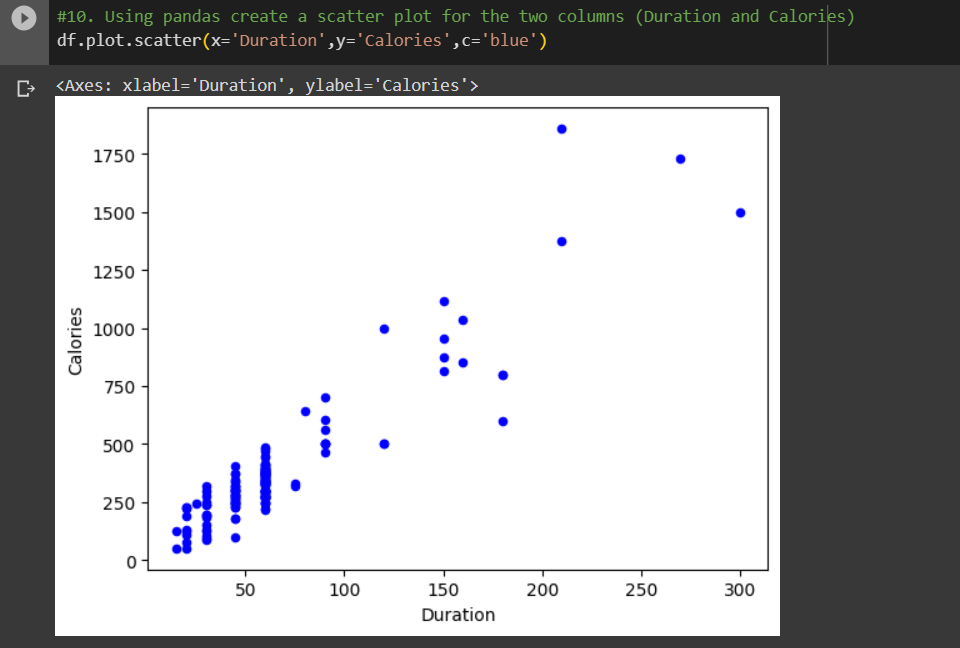
8. Delete the “Maxpulse” column from the main df dataframe



9. Convert the datatype of Calories column to int datatype



10. Using pandas create a scatter plot for the two columns (Duration and Calories)



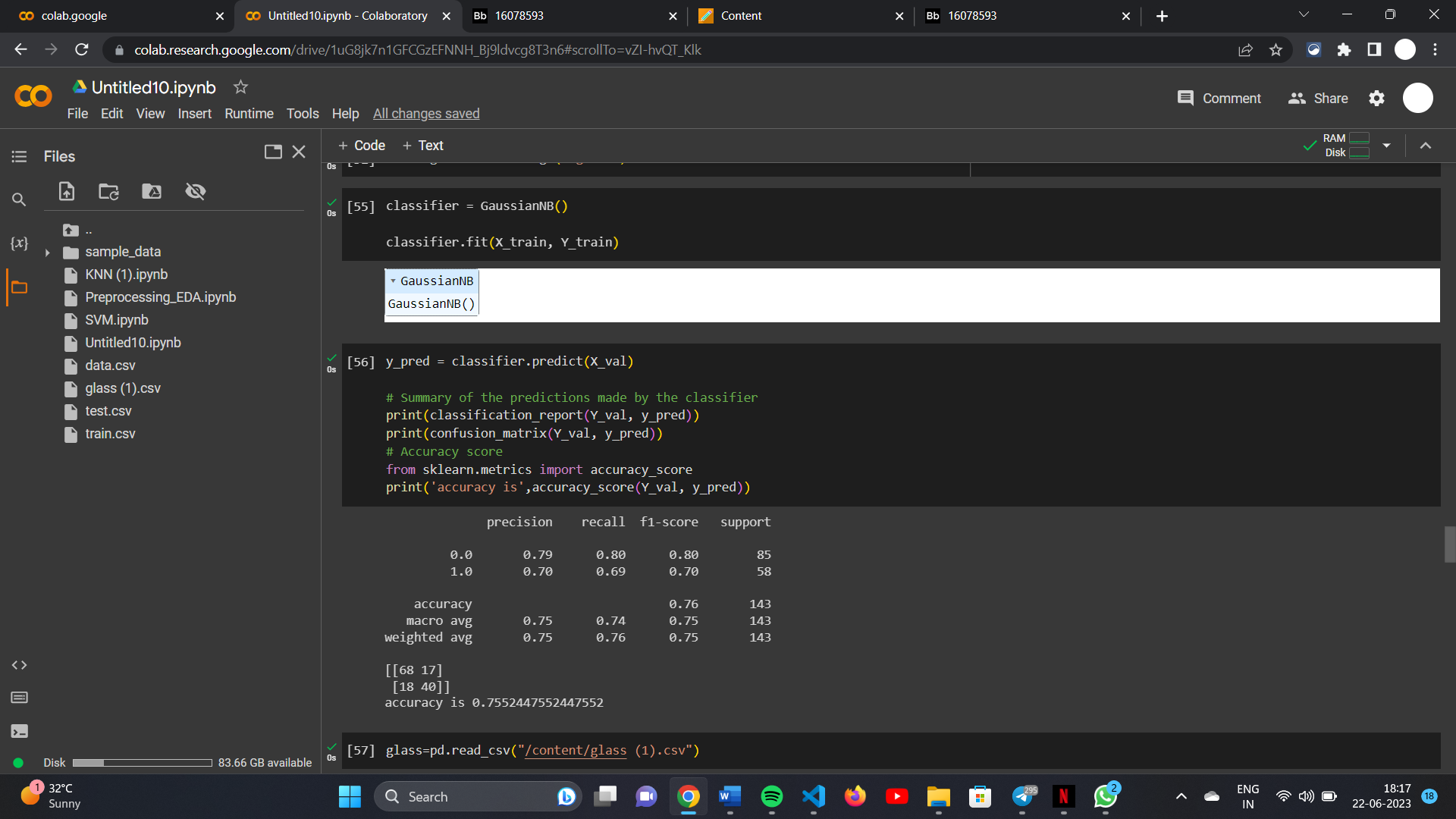
**2. Scikit-learn**

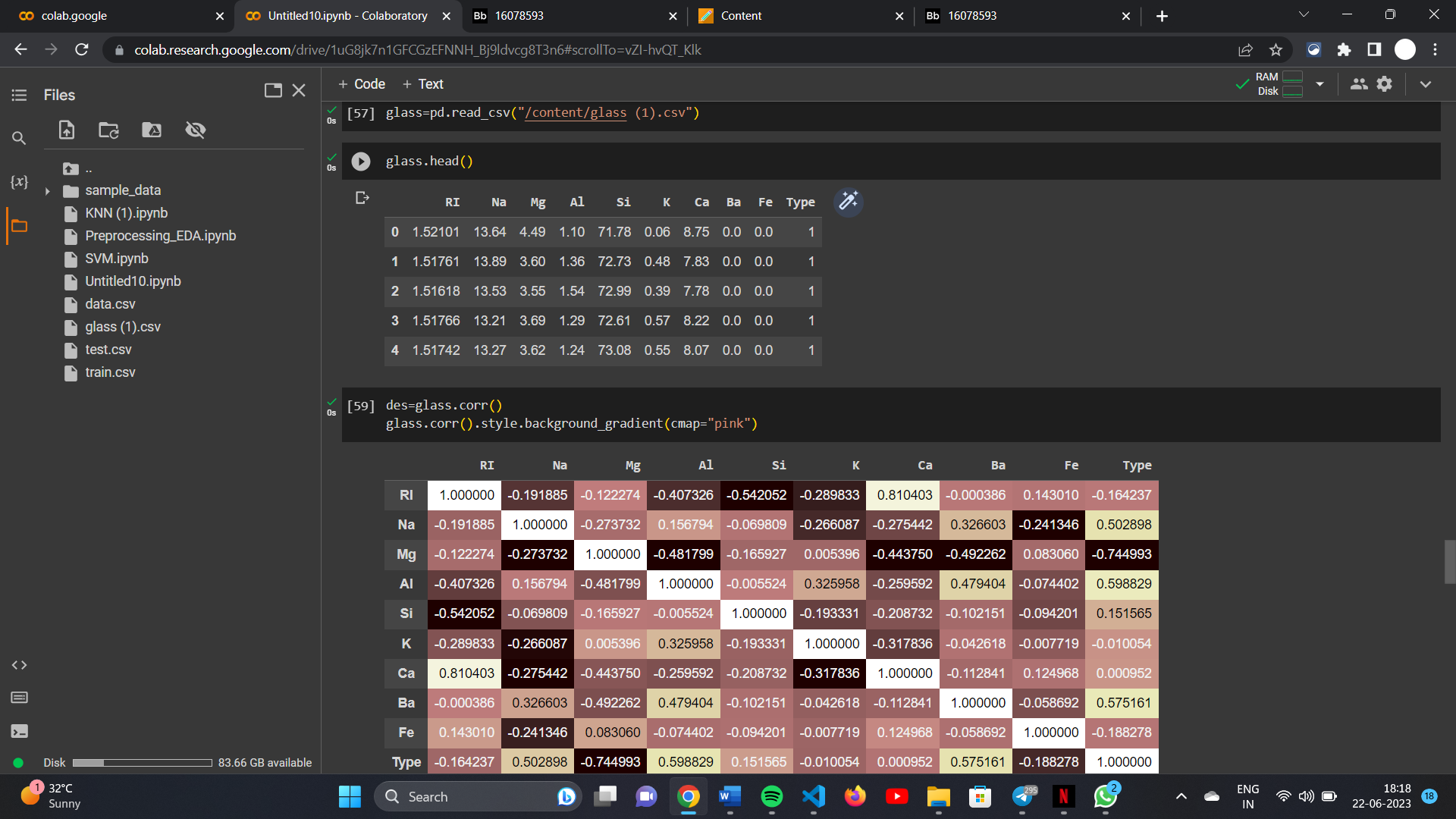
1. Implement Naïve Bayes method using scikit-learnlibrary.

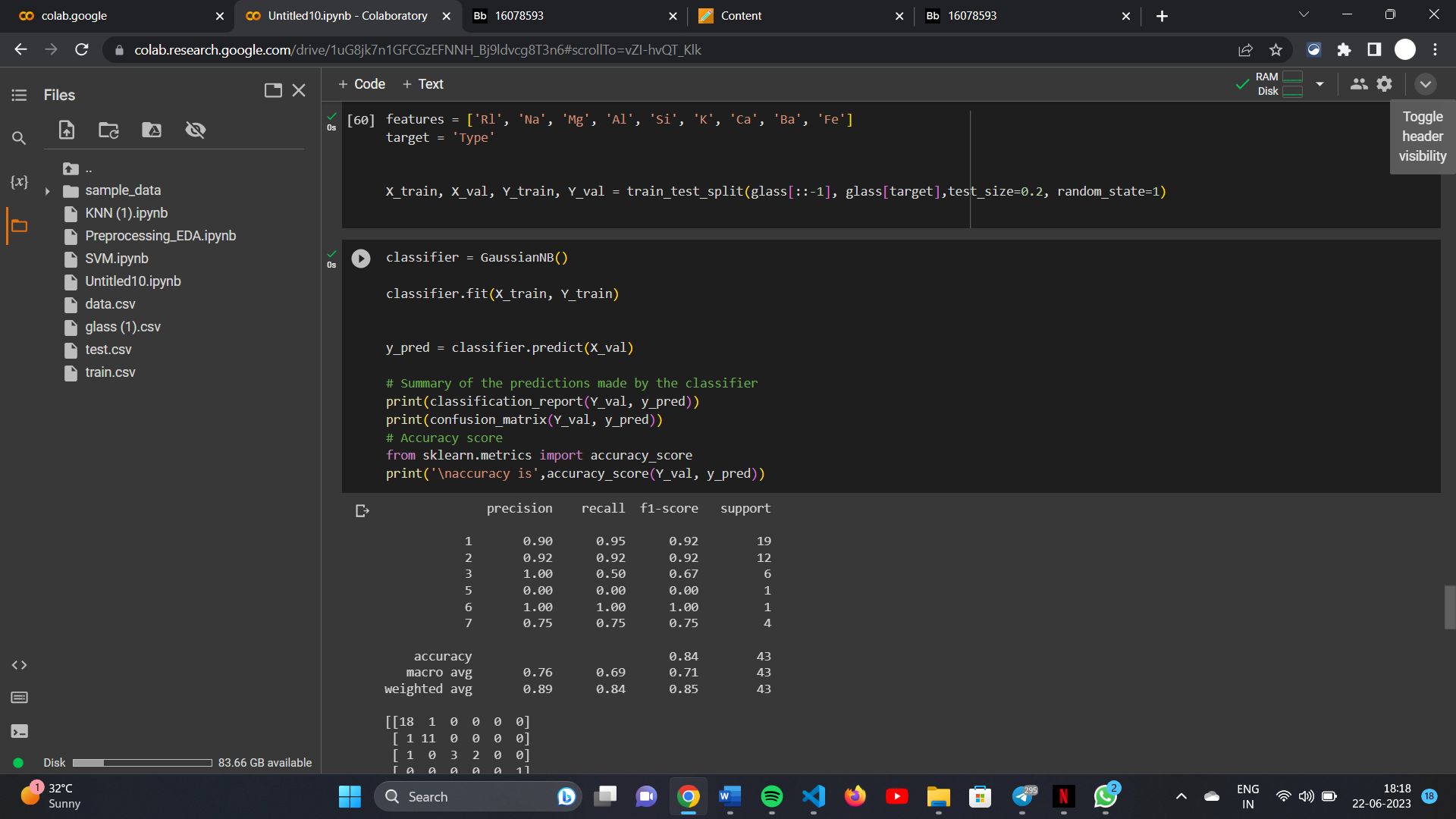
a. Use the glass dataset available in Link also provided in your assignment.

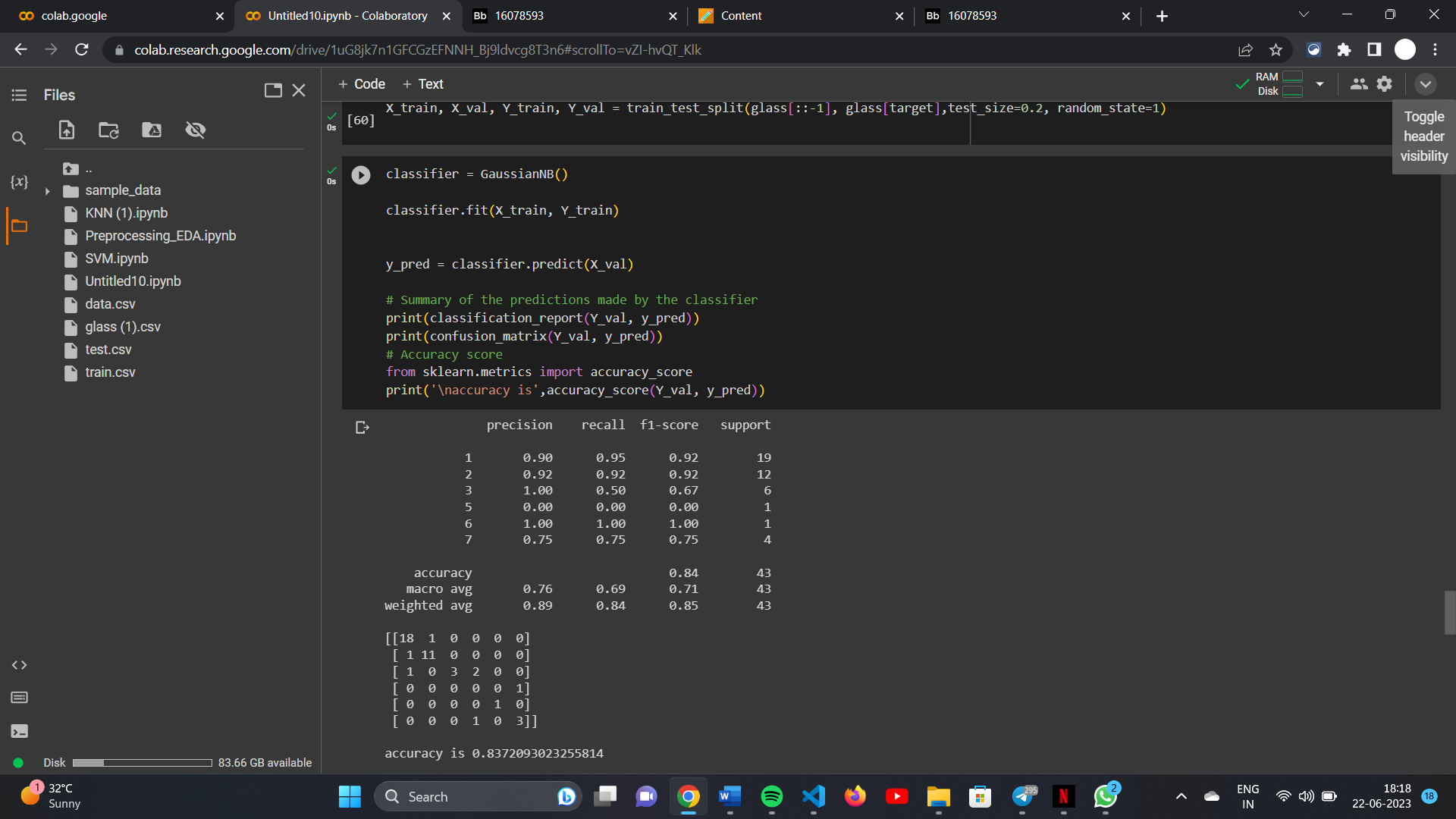
b. Use train\_test\_split to create training and testing part.

2. Evaluate the model on testing part using score and





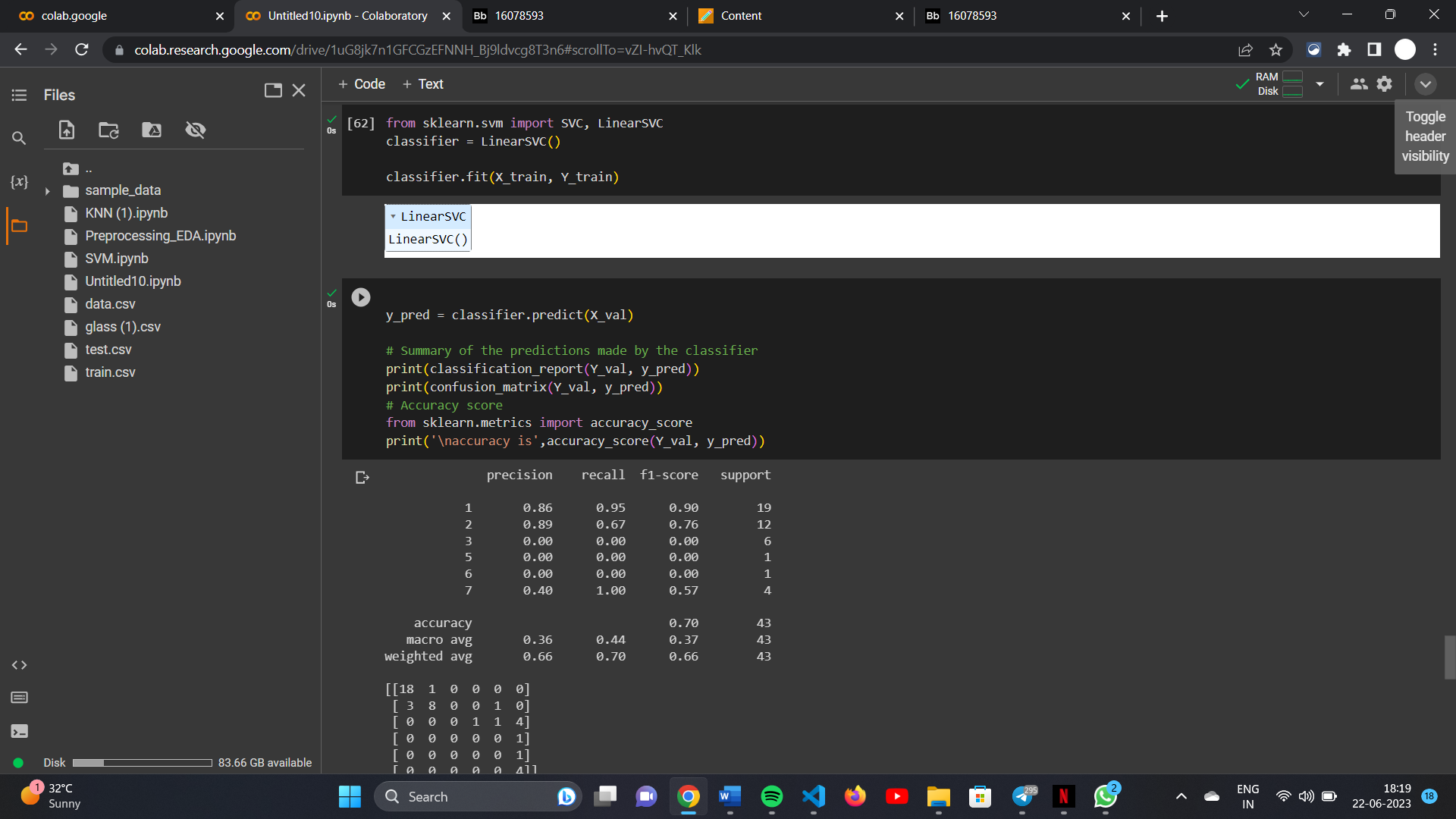


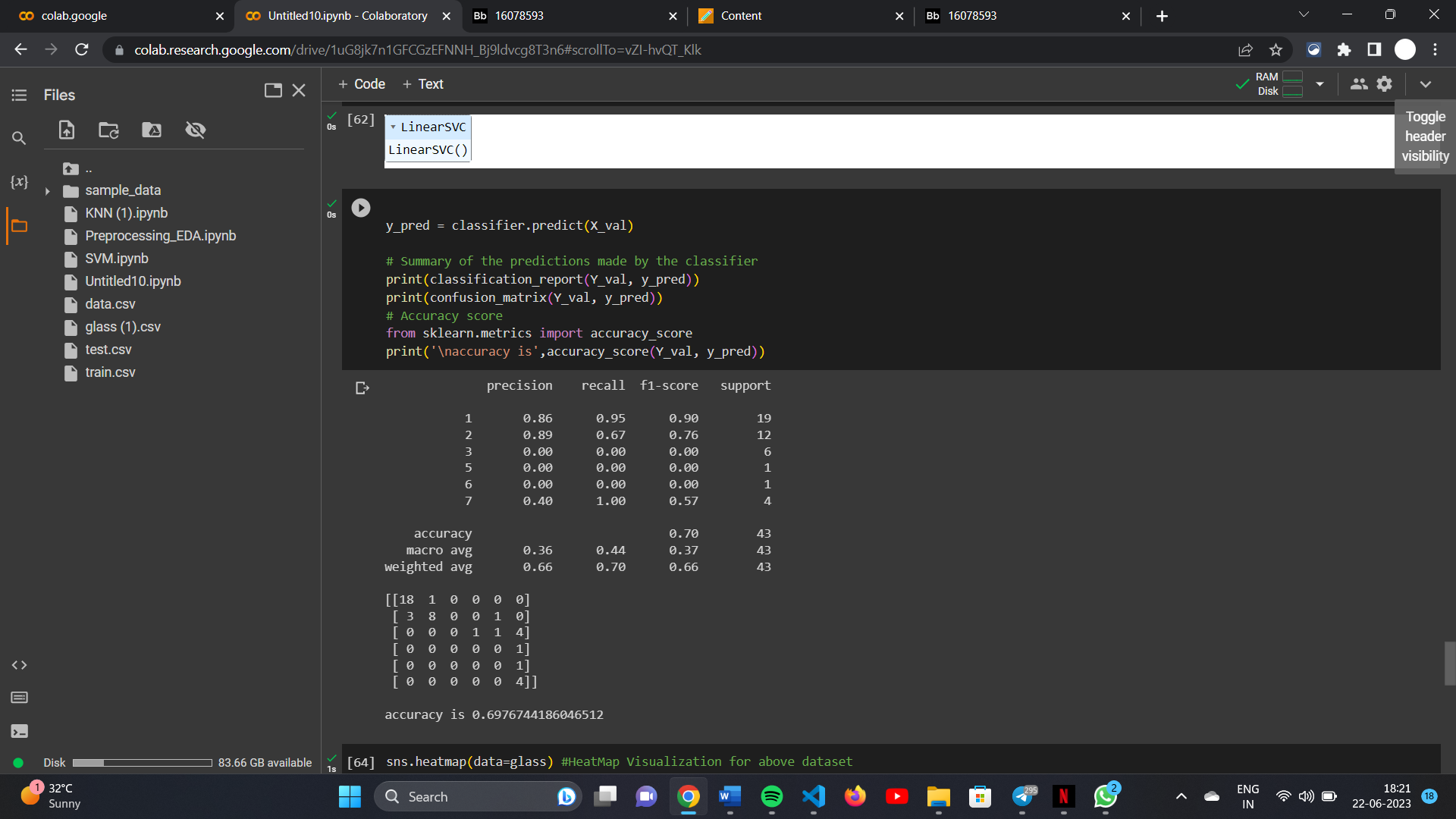


1. Implement linear SVM method using scikit library

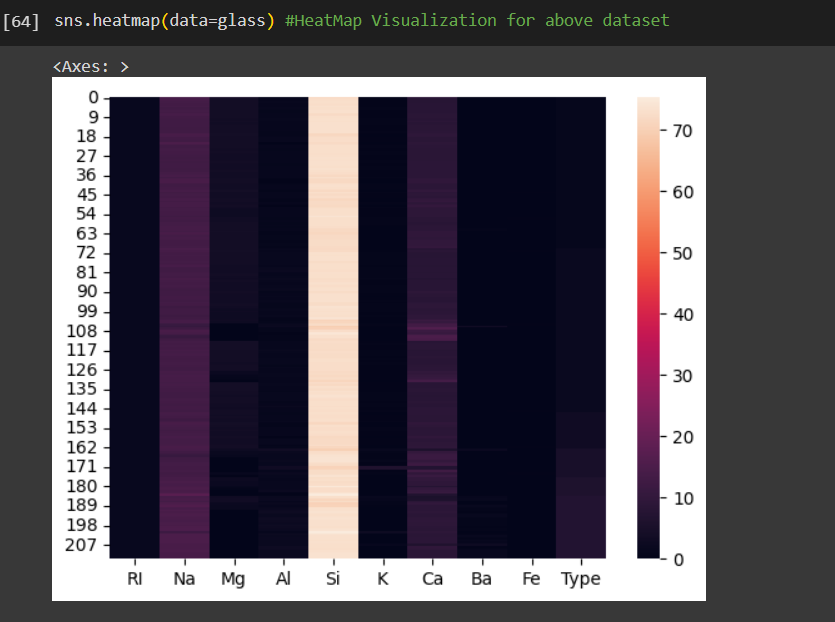
a. Use the glass dataset available in Link also provided in your assignment.

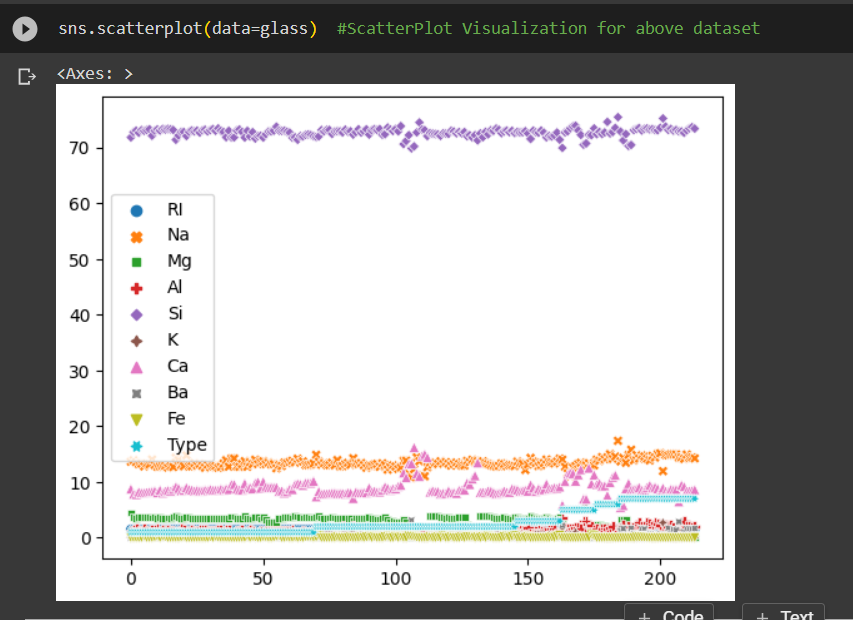
b. Use train\_test\_split to create training and testing part. 2. Evaluate the model on testing part using score and





Do at least two visualizations to describe or show correlations in the Glass Dataset





* Naïve Bayes classifier got the better accuracy
* Naïve Bayes classifier gives better accuracy because it is fast and space efficient
* It is Not sensitive to irrelevant features.
* On the other side SVM is not efficient on large data.